On the Existence of Optimal Stochastic Controls[†]

by

Harold J. Kushner[‡]

1. Introduction.

We will prove several theorems which state, under their prescribed conditions, that if there exists one stochastic control which accomplishes a given task, then there is an optimal stochastic control. The systems of concern are governed by the stochastic vector differential equations

$$dx(\omega, t) = f(x(\omega, t), u(\omega, t))dt + \sigma(x(\omega, t), u(\omega, t))dz(\omega, t) \tag{1}$$

or

$$dx(\omega, t) = f(x(\omega, t), u(\omega, t))dt + dz(\omega, t) \tag{2}$$

where $x(\omega, t)$ is an r-dimensional vector, $u(\omega, t)$ is a vector control, $\sigma(x, u)$ is an $r \times r$ matrix, and $z(\cdot, \cdot)$ is a vector stochastic process. In the form (1), $z(\cdot, \cdot)$ is assumed to be Brownian motion; in the form (2), $z(\cdot, \cdot)$ is a more general process to be described later. Many stochastic systems may be put into the form of (1) or (2), but this will not be pursued here. The problem will be investigated with two types of tasks, or terminal conditions. The first is that $x(\omega, t)$ satisfy (with probability one)

[†]Center for Control Theory, RIAS.

[‡]This research was supported in part by the United States Air Force through the Air Force Office of Scientific Research, Office of Aerospace Research, under Contract No. AF 49(638)-1206 and in part by the National Aeronautics and Space Administration under Contract No. NASw-845.

$$g(x(\omega, T_\omega)) = 0 \tag{3}$$

where g is continuous, for some random terminal time T_ω. The second is a terminal condition on the expectation

$$g(Ex(\omega, T)) = 0, \tag{4}$$

where T is a non-random terminal time. (See [1] for examples where the latter case is of importance.) In both cases, the risk to be minimized by the optimal control is the expectation

$$R(u) = E\int_0^{T_\omega} f_0(x(\omega, t), u(\omega, t)dt \equiv Ex_0(\omega, T_\omega), \tag{5}$$

where T_ω is either the first (random) time that (3) is satisfied or, for the second problem, $T_\omega = T$, any time that (4) is satisfied.

For the deterministic problem, which has already received substantial attention [2], [3], [4], the question of existence (assuming that there is one control which accomplishes the desired task) is equivalent to the question of the closure of a set of attainable states $x(t)$, over all possible controls $u(\cdot)$. In the stochastic case, the question also reduces to that of closure of an appropriate set of attainable sample states or of expectations.

In a well-formulated control problem, not only must the system (1) or (2), the risk (5), and the target (3) or (4) be given, but also the <u>type of observations or information that is to be available to the controller must be given as part of the problem statement</u>. The control can be represented as an explicit function of the observations; e.g., of the form $u(x, t)$, if some components of $x(\omega, t)$ are observed.

In more general cases, values of $x(\omega, s)$ and $z(\omega, s)$ at several instantes of time are available at time t. The control will be a function or functional of these observations. In general, the control may be written as $u(\omega, t)$. The important point is that the optimality of a control refers to optimality within a specific class of control functions†.

In addition to naming the observations upon which the control is to be based, the stopping rule must also be given; the stopping time must also be determinable in terms of the available observations only. With target (3) it is assumed that the arguments of $g(\cdot)$ are observable as well as, perhaps, other information, and that the stopping time is the <u>first time</u> that (3) is satisfied. In general, the stopping time may have a more complicated dependence on ω, or the past history of $x(\omega, s)$, provided that (3) still holds.

2. <u>Existence Theorems</u>.

Define the norm of any vector v with components v_i as $\|v\| = \sum_i |v_i|$. Let K and K_1 be any positive, finite and non-random numbers. The abbreviations a.a.ω and w.p.1 are used for 'almost all ω' and 'with probability one', respectively. Ω is the space of points ω. $\Sigma(t)$ is the minimal σ-field over which $z(\cdot, \tau)$, $\tau \leq t$, is measurable. Define $\tilde{\Sigma}(T) = \Sigma(T) \times \mathcal{J}(T)$, where $\mathcal{J}(T)$ is the Borel field over the interval $[0, T]$. Since $z(\cdot, \cdot)$ is assumed to be measurable in the pair (ω, t) (see (A4)), it is measurable with respect to $\tilde{\Sigma}(T)$, for $t \leq T$. The measure on the sets of $\Sigma(t)$ is $m(d\omega)$ for all t, and the measure on the sets of $\tilde{\Sigma}(T)$ is $\mu(d\omega \times dt)$ for all T.

---

† The distinction will not be pursued further, except for the examples of the theorems. See also discussion in [1], [5] and [6], where we have attempted a beginning of a stochastic variational theory, and have derived necessary conditions for optimality (stochastic Euler equations).

In the proofs, various subsequences of control functions are determined, each is indexed by n, and each is a subsequence of the subsequence determined previous to it.

It will be assumed throught that

(A1) $$|f_i(x, u)| \le K(1 + \|x\| + \|u\|)$$

(A2) $$|f_i(x + \delta x, u + \delta u) - f_i(x, u)| \le K(\|\delta x\| + \|\delta u\|)$$

(A3) $$E \max_{t \le T} \|z(\omega, t)\| < \infty, \quad Ez(\omega, t) = 0.$$

(A4) $z(\cdot,\cdot)$ is measurable in the pair (ω, t), for almost all ω.

(A5) $$|\sigma_{ij}(x, u)| \le K(1 + \|x\| + \|u\|)$$

(A6) $$|\sigma_{ij}(x + \delta x, u + \delta u) - \sigma_{ij}(x, u)| \le K(\|\delta x\| + \|\delta u\|).$$

(A7) $g(\cdot)$ is continuous.

<u>Theorem</u> 1. <u>Assume form</u> (2) <u>and</u> (A1) <u>to</u> (A4) <u>and</u> (A7), <u>and that almost all sample functions of</u> $z(\cdot, \cdot)$ <u>are continuous</u>. <u>Let the family of admissible controls be of the form</u> $u(x, t)$, <u>where</u> $u(\cdot, \cdot)$ <u>satisfies the uniform Lipschitz condition</u>

$$\|u(x + \delta x, t + \delta t) - u(x, t)\| \le K_1(\|\delta x\| + \|\delta t\|) \tag{6}$$

$$\|u(0, 0)\| \le K_1. \tag{7}$$

<u>Let the stopping time be the first random time that</u> (3) <u>holds</u>. <u>Let</u>

$$f_o(x, u) \geq 0 \tag{8}$$

<u>and let the family of admissible controls be further restricted so that, for any admissible control,</u>

$$ET_\omega \leq \hat{T} < \infty \tag{9}$$

<u>where</u> $\hat{T}$ <u>is non-random</u>. <u>Then, if there is one control satisfying</u> (6) <u>such that</u> (3), (8) <u>and</u> (9) <u>are satisfied</u>, w.p.1, <u>then there is an</u> (<u>optimal</u>) <u>control which absolutely minimizes</u> (5) <u>and such that</u> (3) <u>and</u> (9) <u>hold</u> w.p.1.

<u>Proof</u>. Define

$$\alpha = \inf R(u), \tag{10}$$

where the infimum is over all admissible $u(\cdot, \cdot)$. Thus, there is an infinite sequence of controls $u^n(\cdot, \cdot)$ such that $R(u^n) = \alpha_n \to \alpha$ monotonically. To prove the theorem, it must be shown that there is an admissible $\bar{u}(\cdot, \cdot)$ such that $R(\bar{u}) = \alpha$, and (3), (8) and (9) are satisfied w.p.1.[†]

Note first that, under (A1) to (A4), (6) and the continuity of $z(\omega, \cdot)$, for a.a.ω, there is a unique continuous solution to (2), for a.a.ω, such that

$$E \max_{t \leq T} \|x(\omega, t)\| \leq K_2 < \infty, \tag{11}$$

for any finite T. Owing to (6), K_2 does not depend upon the particular $u(\cdot, \cdot)$. (See, for example, the remark on p. 286 of Doob [7], and observe

† The proof is more difficult than the deterministic proof [2], since neither T_ω^n nor $x^n(\omega, t)$ are uniformly bounded in n and ω here.

that Doob's conditions on f (his m) are satisfied uniformly in u here, by (6)).

Extend the domain of definition of each $u^n(\cdot, \cdot)$, if necessary, so that it is defined on all bounded sets. By Ascoli's Theorem, and (6), there is a subsequence of the $u^n(\cdot, \cdot)$ and a $\bar{u}(\cdot, \cdot)$ satisfying (6) so that

$$u^n(x, t) \to \bar{u}(x, t) \tag{12}$$

uniformly on all bounded sets. Let $\bar{x}(\cdot, \cdot)$ correspond to $\bar{u}(\cdot, \cdot)$. It will be shown that $\bar{u}(\cdot, \cdot)$ is an optimal control.

By (A2),

$$\begin{aligned} \delta x^n(\omega, t) \equiv x^n(\omega, t) - \bar{x}(\omega, t) = \int_0^t [f(x^n(\omega, s), u^n(x^n(\omega, s), s)) - \\ f(\bar{x}(\omega, s), \bar{u}(\bar{x}(\omega, s), s))]ds \end{aligned} \tag{13}$$

$$\|\delta x^n(\omega, t)\| \le \int_0^t Kr(\|\delta x^n(\omega, s)\| + \|u^n(x^n(\omega, s), s) - \bar{u}(\bar{x}(\omega, s), s)\|)ds.$$

Substituting

$$\begin{aligned} &\|u^n(x^n, s) - \bar{u}(\bar{x}, s)\| \le \\ &\qquad \|u^n(x^n, s) - u^n(\bar{x}, s)\| + \|u^n(\bar{x}, s) - \bar{u}(\bar{x}, s)\| \end{aligned}$$

and

$$\|u^n(x^n, s) - u^n(\bar{x}, s)\| \le K_1\|\delta x^n\|$$

into (13) yields

$$\|\delta x^n(\omega, t)\| \leq K_2 \int_0^t (\|\delta x^n(\omega, s)\| + \delta u^n(\omega, s))ds$$

where

$$\delta u^n(\omega, s) \equiv \|u^n(\bar{x}(\omega, s), s) - \bar{u}(\bar{x}(\omega, s), s)\|$$

from which follows

$$\max_{t \leq T} \|\delta x^n(\omega, t)\| \leq K_3 \int_0^T \delta u^n(\omega, s)ds \tag{14}$$

for any finite T, where K_3 depends on T.

Define $A(p, T) = \{\omega, t: \|\bar{x}(\omega, t)\| \leq p, \ t \leq T\}$ and $\bar{A}(p, T)$ as the complement of $A(p, T)$ where T is held fixed. From (14),

$$E \max_{t \leq T} \|\delta x^n(\omega, t)\| \leq R_1 + R_2,$$

where

$$\begin{aligned} R_1 &= K_3 \int_{A(p,T)} \delta u^n(\omega, t)\mu(d\omega \times dt) \\ R_2 &= K_3 \int_{\bar{A}(p,T)} \delta u^n(\omega, t)\mu(d\omega \times dt). \end{aligned} \tag{15}$$

Equation (11) implies that

$$\text{Prob.} \{ \max_{s \leq T} \|x(\omega, s)\| > p\} \to 0$$

uniformly in the control, as $p \to \infty$. Thus $\mu(\bar{A}(p, T)) \to 0$ as $p \to \infty$. Since, if T is finite, R_2 exists and is finite for any $\bar{A}(p, T)$, it tends to zero as $\mu(\bar{A}(p, T)) \to 0$. Thus, for any $\epsilon > 0$, there is a p_ϵ such that, for $p \geq p_\epsilon$, we have $R_2 \leq \epsilon$. Since $u^n(x, t) \to \bar{u}(x, t)$ uniformly on all bounded sets, for any finite $p > 0$, $T > 0$ and $\epsilon > 0$, there is an $N_\epsilon(p, T)$ such that $\delta u^n(\omega, t) \leq \epsilon/TK_3$ for $n \geq N_\epsilon(p, T)$ and for (ω, t) in $A(p, T)$. Thus, for large n,

$$E \max_{t \leq T} \|\delta x^n(\omega, t)\| \leq 2\epsilon.$$

It is now obvious that there is a subsequence such that

$$E \max_{t \leq T} \|\delta x^n(\omega, t)\| \to 0 \tag{16}$$

w.p.1. (Take, for example, a sequence of $\epsilon \to 0$; the subsequence indexed by $N_\epsilon(p_\epsilon, T)$ satisfies (16).) Thus there is a further subsequence such that, for finite T and w.p.1,

$$\max_{t \leq T} \|\delta x^n(\omega, t)\| \to 0. \tag{17}$$

It will be shown next that (3) is satisfied with probability $\geq 1-\epsilon$, for any $\epsilon > 0$. Define, for a.a.ω,

$$T_\omega = \liminf_n T_\omega^n. \tag{18}$$

T_ω is also a random time (i.e., does not depend on the future) since the $T_\omega^n(T)$ are random times and, by (18),

$$\{\omega: T_\omega \le t\} \in \Sigma(t).$$

Also, by (9),

$$ET_\omega \le \hat{T} < \infty.$$

Let $B_\epsilon(\delta) = \{\omega: T_\omega \ge T - \delta\}$, where $\delta > 0$. Thus, for any $\epsilon > 0$, there is a T such that

$$m(B_\epsilon(\delta)) < \epsilon.$$

We have

$$S = |g(x^n(\omega, T_\omega^{\ n})) - g(\bar{x}(\omega, T_\omega))| \le S_1 + S_2$$

$$S_1 = |g(x^n(\omega, T_\omega^{\ n})) - g(\bar{x}(\omega, T_\omega^{\ n}))|$$

$$S_2 = |g(\bar{x}(\omega, T_\omega^{\ n})) - g(\bar{x}(\omega, T_\omega))|.$$

Since $g(x^n(\omega, T_\omega^{\ n})) = 0$ w.p.1, S depends on ω but not on n. We will now evaluate S for each fixed, ω in $B_\epsilon(\delta)$.

$$\lim_n S \ge \lim_n \inf S_1 + \lim_n \inf S_2.$$

Now <u>fix</u> ω <u>in</u> $\Omega - B_\epsilon(\delta)$. By (18), there is a subsequence $T_\omega^{\ n} \to T_\omega$. Also, $|T_\omega^{\ n} - T| < \delta$ except for a finite number of terms. Thus, by the continuity of $g(\cdot)$ and of $\bar{x}(\omega, \cdot)$, for $t \le T$, $\lim_n \inf S_2$ is zero. By the uniform convergence implied by (17), $\lim_n \inf S_1$ is also zero. Thus (3) is satisfied with probability greater than $1 - \epsilon$. Since ϵ is arbitrary, (3) is satisfied w.p.1.

We now show that $R(\overline{u}) = \alpha$ w.p.1., or that

$$\lim_n Q_n = 0$$

$$Q_n = Ex_o^n(\omega, T_\omega^n) - E\overline{x}_o(\omega, T_\omega)$$

Since $\lim_n R(u^n) = \alpha$ we have $\lim_n Q_n^n \leq 0$. Let

$$Q_n = Q_1^n + Q_2^n$$

$$Q_1^n = Ex_o^n(\omega, T_\omega^n) - E\overline{x}_o(\omega, T_\omega^n)$$

$$Q_2^n = E\overline{x}_o(\omega, T_\omega^n) - E\overline{x}_o(\omega, T_\omega).$$

Take a subsequence so that both $Ex_o^n(\omega, T_\alpha^n)$ and $E\overline{x}_o(\omega, T_\omega^n)$ converge monotonically. Since $\overline{x}_o(\omega, T_\omega^n) \geq 0$, we may apply Fatou's Lemma and obtain

$$\lim_n \int \overline{x}_o(\omega, T_\omega^n)m(d\omega) \geq \int \liminf_n \overline{x}_o(\omega, T_\omega^n)m(d\omega).$$

Since $f_o(x, u) \geq 0$, $\overline{x}_o(\omega, \cdot)$ is nondecreasing in t. It is also continuous at all finite t. Since $\liminf_n T_\omega^n < \infty$ w.p.1, we have $\liminf_n \overline{x}_o(\omega, T_\omega^n) = \overline{x}_o(\omega, T_\omega)$ w.p.1. Thus for sufficiently large n,

$$Q_1^n \geq 0.$$

w.p.1. Now, if we can show that there is a subsequence such that $\lim_n Q_1^n = 0$ w.p.1., we will have proved that $Q_n = 0$ w.p.1.

If T_ω^n were uniformly bounded in n and ω, then $\lim Q_1^n \to 0$ by virtue of (16). Since T_ω^n is not uniformly bounded, we take the following approach. Let $B^n(T) = \{\omega: T_\omega^n > T\}$. By a standard inequality (see [8], p. 157)

$$m(B^n(T)) \le \frac{ET_\omega^n}{T} \le \frac{\hat{T}}{T}.$$

Thus, there is a $T_i < \infty$ such that

$$m(B^n(T_i)) \le \epsilon 2^{-i}$$

for any n. Let δ_i be a sequence of positive numbers tending to zero. By (16) we may choose a subsequence of the $u^n(\cdot,\cdot)$ which we will index by i, so that

$$E \max_{t \le T_i} |x_0^i(\omega, t) - \bar{x}_0(\omega, t)| < \delta_i.$$

Define $B_\epsilon = \bigcup_i B^i(T_i)$, and $\bar{B}_\epsilon$ as the complement of B_ϵ. We have

$$m(B_\epsilon) < \epsilon.$$

When ω is in $\bar{B}_\epsilon$, we have $T_\omega^i \le T_i$ for all i. Also

$$\int_{\bar{B}_\epsilon} |x_0^i(\omega, T_\omega^i) - \bar{x}_0(\omega, T_\omega^i)| m(d\omega) < \delta_i \to 0$$

w.p.1. Thus $\lim Q_1^n = 0$ with probability $\ge 1 - \epsilon$. Since ϵ is arbitrary, $\lim Q_1^n = 0$ w.p.1, and we conclude that $\lim Q_n = 0$ and $R(\bar{u}) = \alpha$ w.p.1.

Since $f_0(x, u) \geq 0$, the sample functions $\tilde{\tilde{x}}_0(\omega, t)$ are non-decreasing functions of time. Let $\tilde{T}_\omega \leq T_\omega$ be the first time that (3) is satisfied. Since α is the minimum possible risk, and it is attained by stopping at T_ω, it will also be attained by stopping at $\tilde{T}_\omega$. Thus, we have proved that the first time (3) is satisfied is optimum, and the proof is concluded.

An easy consequence of the proof is a corresponding result when the terminal constraint is a set of expectations. Note that $f_0 \geq 0$ and the continuity of almost all $z(\omega, \cdot)$ are not required.

Corollary. Assume (A1) to (A4) and (A7). Let all admissible controls satisfy (6) and (7). Let the target set be the set of expectations such that (4) is satisfied, where T is non-random and finite. Then if there is one admissible control such that (4) is satisfied, then there is an admissible optimal control.

We state the following theorem for form (1). The proof, although differing in detail, is essentially the same as the proof of Theorem 1, and will not be given. $z(\cdot, \cdot)$ is confined to Brownian motion to assure that the various stochastic integrals exist and have suitable properties.

Theorem 2. Assume all the conditions of Theorem 1, except let (1) replace (2), and let $z(\cdot, \cdot)$ be vector Brownian motion. Assume (A5) and (A6). Then, the conclusions of Theorem 1 hold.

The existence and uniqueness of solutions to (1) and (2) has not yet been proved under much more general conditions on $u(x, t)$ than those of Theorem 1. In this sense, Theorems 1 and 2 represent about the best attainable result with the use of the control form $u(x, t)$.

In order to present existence results with other control forms, a different approach is taken in the sequel. Assume that the information upon which the values of the control are to depend are observations on $z(\cdot, \cdot)$, and that almost all sample functions of these observations are Lebesgue measurable. Without specifying the type of observations further, it can be said that there is a sub σ-field $\tilde{\Sigma}_c(t) \subseteq \tilde{\Sigma}(t)$ which is the minimal σ-field with respect to which the observations at time t are measurable. There is also a sub σ-field $\tilde{\Sigma}_c(T) \subseteq \Sigma(T)$ with sections $\Sigma_c(t)$ and $\mathcal{J}(T)$ for almost all $t \le T$ and ω, respectively, and with respect to which, the observations, as functions of ω and t, are measurable.

A further condition condition that we impose is that any time interval $[0, T]$ can be divided into half open intervals such that, if s and t are in the same interval and $s < t$, then $\Sigma_c(s) \subseteq \Sigma_c(t)$. The meaning of this condition is simply that, within each interval, the information available to the controller does not decrease with time. That is to say, that the values of all observations that are available at s are also available at t. At certain discrete times, however, information may be lost, if desired; i.e., the memory may overflow or old information may be replaced. Let $t_o < t \le t_1$ be one of the intervals. Then the observations at time s, $t_o < s \le t$, are measurable with respect to $\Sigma_c(t)$. There are cases where these restrictions may be avoided, e.g., when $u(\cdot, \cdot)$ is a scalar values control, but this will not be pursued.

The control form $u(\omega, t)$ will be used. It is obvious that:

(A8) All admissible $u(\cdot, t)$ and $u(\cdot, \cdot)$ are measurable with respect to $\Sigma_c(t)$ and $\tilde{\Sigma}_c(T)$, for any finite T and t. Also all $u(\omega, \cdot)$ are Lebesgue measurable.

We will also require

(A9) Let $u(\omega, t)$ take values only in the convex compact set U. Let $k(\cdot)$ be a continuous function and $k(U)$ a convex set.

The following theorems are also of interest for the reason that if $\tilde{\Sigma}_c(T) = \tilde{\Sigma}(T)$ and $\Sigma_c(t) = \Sigma(t)$, then the optimal control in this family of controls yields at least as small a risk as the optimal control of any other family. There are a number of important cases where the control may be chosen based on observations on $z(\cdot, \cdot)$ only; e.g., if $dz(\cdot, \cdot) = 0$, except at random times determined by a Poisson or other distribution, when it takes an impulsive form; or when a stochastic process, say $z_0(\cdot, \cdot)$, correlated with the $z(\cdot, \cdot)$ which drives the dynamical system, is the only function which is observed.

The following Lemma will be useful.

Lemma. Let the interval $[0, T]$ be divided into finitely many half open intervals. Let $\Sigma_c(s) \subsetneq \Sigma_c(t)$ if $s < t$ and both are in the same interval. Let $k(\cdot)$ be continuous and the bounded vector valued function $\gamma(\cdot, \cdot)$ measurable with respect to $\tilde{\Sigma}_c(T)$. If there exists one, not necessarily measurable, function $u(\cdot, \cdot)$ with values in the compact set U and such that

$$\gamma(\omega, t) = k(u(\omega, t))$$

almost everywhere, then $u(\cdot, t)$ may be defined to be measurable with respect to $\Sigma_c(t)$.

Proof. We may confine our attention to a single interval. Under the hypothesis on $k(\cdot)$, $u(\cdot, \cdot)$, U and $\gamma(\cdot, \cdot)$, a theorem of Wazewski [9] yields the existence of a Lebesgue measurable $u(\omega, \cdot)$ for each fixed ω, such that $\gamma(\omega, t) = k(u(\omega, t))$. If the observations take the same values for ω_1 and ω_2 provided $s \le t$, then ω_1 and ω_2 are always in the same sets of $\Sigma_c(s)$, and $\gamma(\omega_1, s) = \gamma(\omega_2, s)$ for $s \le t$. Hence, by the method of construction in [9], $u(\omega_1, s) = u(\omega_2, s)$, $s \le t$. Thus this $u(\omega, t)$ does not depend upon the future values of the observations and, hence, at t, is measurable with respect to $\Sigma_c(t)$.

Theorem 3. Assume (A1) to (A4), (A7) and (A8). Let

$$dx(\omega,t) = A(t)x(\omega,t)dt + k(u(\omega,t))dt + dz(\omega,t), \tag{19}$$

$$dx_o(\omega,t) = \sum_{j=1}^{r-1} a_{oj}x_j(\omega,t)dt + k_o(u(\omega,t))dt,$$

where $A(\cdot)$ is bounded and Lebesgue measurable and $k(\cdot)$ satisfies (A9). Let the terminal condition be the set (4) of expectations,† and let the terminal time of control T be non-random and, for all admissible controls let

$$T \le \hat{T} < \infty. \tag{20}$$

Then, if there exists one admissible control such that (4) holds w.p.1. and (19) holds. Then there is an optimal control such that (4) holds w.p.1. and (19) holds.

Proof. Under (A1) to (A4) and (A8) and (A9), the existence and uniqueness w.p.1. of the solution to (19) for a.a.ω is a special case of the theorem in the appendix of [6]. Again, let $\alpha = \inf R(u)$, where the infi-um is over the class of controls of the hypothesis of this theorem. Let $u^n(\cdot, \cdot)$ be a sequence of controls such that (4) is satisfied at $T^n \le \hat{T}$ w.p.1., and $R(u^n) = \alpha^n$ monotonically decreases to α. We must show that there is a $\bar{u}(\cdot, \cdot)$ and a T such that $R(\bar{u}) = \alpha$ and $g(E\bar{x}(\omega, T)) = 0$ w.p.1.

† See remark at the end of the proof.

Let $\Phi(s, t)$ be the fundamental matrix of the system (19). By the hypothesis on $A(\cdot)$, $\Phi(\cdot, \cdot)$ is finite, Lebesgue integrable, and continuous in both arguments. We have

$$Ex^n(\omega, T^n) = E\int_0^{T^n} \Phi(T^n, t)k(u^n(\omega,t))dt + \Phi(T^n, 0)x(0)$$
$$+ E\int_0^{T^n} \Phi(T^n, t)dz(\omega, t). \tag{21}$$

Since $\Phi(s, t)$ does not depend on the values of $z(\cdot, \cdot)$, there is no trouble in defining the last integral of (21). Since $Ez(\omega, t) \equiv 0$, the last integral is zero for any finite upper limit of integration. Since $g(\cdot)$ is continuous, and the value $u^n(\omega, t)$ is confined to compact U, and each $k(u^n(\cdot, \cdot))$ is measurable with respect to $\tilde{\Sigma}_c(T)$, we have

$$\int_0^T E\|k(u^n(\omega, t))\|dt \le K_1$$

for some K_1 not depending on n, where we define

$$T = \lim_n \inf T^n \le \hat{T} < \infty.$$

Thus, there is a subsequence, and a function $\gamma(\cdot, \cdot)$ measurable with respect to $\tilde{\Sigma}_c(T)$, such that

$$k(u^n(\cdot, \cdot)) \xrightarrow{w} \gamma(\cdot, \cdot)$$

where by <u>weak convergence</u> we mean

$$\int_A k(u^n(\omega, t))\mu(d\omega \times dt) \to \int_A \gamma(\omega, t)\mu(d\omega \times dt), \tag{22}$$

for any set A in $\tilde{\Sigma}_c(T)$.

Next, following the method of Roxin [4], it will be shown that, for a.a. fixed ω and t, there is a number $u(\omega, t)$ in U such that

$$\gamma(\omega, t) = k(u(\omega, t)). \tag{23}$$

For any set A in $\tilde{\Sigma}_c(T)$, (22) and the uniform boundedness of $\|k(u^n(\omega, t))\|$ and Fatou's Lemma, yield, for any vector y,

$$\int_A \operatorname*{lub}_{v \in U} y'k(v(\omega, t))\mu(d\omega \times dt) \geq \int_A y'\gamma(t)\mu(d\omega \times dt).$$

$$\geq \lim_n \int_A y'k(u^n(\omega,t))\mu(d\omega \times dt) = \int_A y'\gamma(\omega,t)\mu(d\omega \times dt).$$

Since the reverse inequalities hold with glb and lim inf replacing lub and lim sup, respectively, we have, for a.a. fixed ω and t,

$$\operatorname*{lub}_{v \in U} y'(k(v(\omega, t)) \geq y'\gamma(\omega, t) \geq \operatorname*{glb}_{v \in U} y'k(v(\omega, t)). \tag{24}$$

Redefine $\gamma(\cdot, \cdot)$ on the remaining set of measure zero so that (24) holds everywhere.

Since U, the range of $v(\omega, t)$ in (23), is a closed convex set, (24) implies that, for each ω and t, there is a $u(\omega, t)$ in U such that (23) holds.

According to the lemma, $u(\cdot, \cdot)$ can be defined to be measurable with respect to $\tilde{\Sigma}_c(T)$.

It will now be shown that $\bar{u}(\cdot, \cdot) = u(\cdot, \cdot)$ and T are the optimal control and stopping time, respectivefully.

Let n index a subsequence such that T^n tends monotonically to T. Letting $\bar{x}(\cdot, \cdot)$ correspond to $\bar{u}(\cdot, \cdot)$, we have

$$\bar{x}(\omega, t) = \Phi(t,0)x(0) + \int_0^t \Phi(t,s)k(\bar{u}(\omega,t))ds + \int_0^t \Phi(t,s)dz(\omega,t). \tag{25}$$

Equation (25) has a unique solution with finite mean w.p.1. Thus, w.p.1.,

$$\begin{aligned} \|Ex^n(\omega,T^n) - E\bar{x}(\omega,T)\| \le \|E\int_0^T \Phi(T,t)\Big[k(u^n(\omega,t)) - k(\bar{u}(\omega,t))\Big]dt\| \\ + \|E\int_T^{T^n} \Phi(T^n,t)k(u^n(\omega,t))dt\| + \|E\int_T^{T^n} \Phi(T^n,t)dz(\omega,t)\|. \end{aligned} \tag{26}$$

The last term on the right of (26) is identically zero. The next to last term converges to zero since the integrals exist and $T^n \to T$. The first term converges to zero by virtue of the weak convergence of $k(u^n(\cdot, \cdot))$ to $k(\bar{u}(\cdot, \cdot))$. Thus

$$Ex^n(\omega, T^n) \to E\bar{x}(\omega, T)$$

w.p.1. Thus, $Ex_o^n(\omega, T^n) = \alpha^n \to E\bar{x}_o(\omega, T) = \alpha$. Since $g(\cdot)$ is continuous $g(F\bar{x}(\omega, T)) = 0$ and the proof is concluded.

Remark. A stronger type of convergence than the weak convergence argument used here appears to be necessary to establish convergence of the sample functions in general (or even of their expectations in the non-linear case) as was done in Theorem 1. This is the reason for restricting the target to a set of expectations, rather than sample functions, and the linear assumption (19). It would be useful to prove whether or not sample function convergence is necessary, in general, in order to satisfy terminal constraints on $x(\omega, t)$.

If $u(\cdot, \cdot)$ is scalar valued, then a sequence $\hat{u}^n(\cdot, \cdot)$ may be constructed so that $\bar{u}(\omega, t) = \liminf_n \hat{u}^n(\omega, t)$ satisfies (23) and is measurable with respect to $\tilde{\Sigma}_c(T)$. (See [4], where equation (2.15) is valid in the scalar case and may be extended to our case.)

In Theorem 4, the assumption that Ω is denumerable allows us to obtain a convergent subsequence of sample functions and apply the terminal constraint $g(x(\omega, T_\omega)) = 0$.

Theorem 4. Assume the hypothesis of Theorem 1 on $z(\cdot, \cdot)$ except let Ω be denumerable. $\Omega = \{1, \ldots, n, \ldots\}$ with $p_i > 0$ the probability that $\omega = i$. Assume (1) in the particular form

$$dx(i, t) = c(x(i, t))dt + h(x(i, t))k(u(i, t)dt + dz(i, t), \qquad (27)$$

where $c(\cdot)$ satisfies (A1) and (A2), and $h(\cdot)$ is an $r \times r$ matrix satisfying (A5) and (A6). Let $k(\cdot)$ satisfy (A9) and let all admissible controls satisfy (A8) (both of Theorem 3). Let the terminal time of control be the first random time that (3) holds and assume (8) and also that (9) holds for any admissible control. Then, if there is one control satisfying (A8) such that (3) and (9) hold, w.p.1. then there is an (optimal) control which absolutely minimizes (5), and is such that (3) and (9) hold w.p.1., and the stopping time is the first time that (3) holds.

Proof. The proof of the existence and uniqueness of the solutions to (27), and the property that the sample functions are continuous and

$$\sum_i p_i \max_{t \le T} \|x(i, t)\| = E \max_{t \le T} \|x(\omega, t)\| < \infty \qquad (28)$$

for all finite T, is an easy consequence of the theorem in the appendix of [6]. Let $u^n(\cdot, \cdot)$ be a sequence of controls such that (3) is satisfied at the first time $T_\omega{}^n$ and $R(u^n) = \alpha^n \to \alpha = \inf R(u)$. We will again find a $\bar{u}(\cdot, \cdot)$ and T_ω such that $R(\bar{u}) = \alpha$ and the first time (3) is satisfied is T_ω. Let $\bar{x}(\cdot, \cdot)$ correspond to $\bar{u}(\cdot, \cdot)$. Let $\delta x^n(\omega, t) = x^n(\omega, t) - \bar{x}(\omega, t)$. In fact, it is only necessary to find a $\bar{u}(\cdot, \cdot)$ and a T_ω such that, for some subsequence,

$$E \max_{t \le T} |\delta x^n(\omega, t)| \to 0 \qquad (29)$$

$$\max_{t \le T} |\delta x^n(\omega, t)| \to 0 \qquad (30)$$

uniformly in n for all finite T w.p.1. The rest of the proof is exactly the same as the corresponding part of the proof of Theorem 1 and will not be given.

For $t > T_\omega^n$, define $u^n(\omega, t)$ in any convenient way that satisfies (A8) and (A9). Since $u^n(\omega, t)$ is in the compact set U, and since $k(\cdot)$ is continuous, $k(u^n(\omega, t))$ is uniformly bounded in n, ω and t. Thus for any finite T there is a subsequence such that

$$k(u^n(\omega,t)) \xrightarrow{w} \gamma(\omega, t),$$

where, by weak convergence we mean

$$\int_A k(u^n(\omega,t))\mu(d\omega \times dt) \xrightarrow{w} \int_A \gamma(\omega, t)\mu(d\omega \times dt) \qquad (31)$$

for any set A in $\tilde{\Sigma}_c(T)$.† $\gamma(\omega, t)$ is thus defined on the open interval $[0, \infty)$ for each ω.

Since the set $A_i = \{(\omega, t)\colon \omega = i,\ t \text{ in } B$, a linear Borel set bounded by $T\}$ is in $\tilde{\Sigma}_c(T)$, and since $\mu(A_i) = p_i \ell(B) > 0$, where $\ell(B)$ is the Lebesgue measure of B, we also have the sample function weak convergence,

$$\int_B k(u^n(\omega, t))dt \to \int_B \gamma(\omega, t)dt \qquad (32)$$

for all bounded linear Borel sets B, and all ω.

† For $t > T_\omega^n$, it is not necessary to define $u^n(\omega, t)$. It is only necessary to satisfy (31) on the restricted space $\{(\omega, t)\colon t \le T_\omega\}$ which has the finite measure ET_ω. This approach, involving more details than the one used above, yields the same result.

The Lemma proceding Theorem 3 proves the existence of a $u(\cdot, \cdot)$ satisfying (A8) and (A9) such that

$$k(\bar{u}(\omega, t)) = \gamma(\omega, t). \tag{33}$$

$\bar{u}(\omega, \cdot)$ is thus defined on all compact time sets and, hence, on the open interval $[0, \infty)$. It will next be shown that $\bar{u}(\cdot, \cdot)$ is the optimum control. We have

$$\delta x^n(\omega, t) \equiv x^n(\omega, t) - \bar{x}(\omega, t) = \int_0^t \Big[c(x^n(\omega, s)) - c(\bar{x}(\omega, s))\Big]ds + \tag{34}$$
$$\int_0^t \Big[h(x^n(\omega, s)) - h(\bar{x}(\omega, s))\Big]k(u^n(\omega, s))ds + q_\omega^n(t),$$

where

$$q_\omega^n(t) = \int_0^t h(\bar{x}(\omega, s))\Big[k(u^n(\omega, s)) - k(\bar{u}(\omega, s))\Big]ds. \tag{35}$$

From (28), and the fact that $p_1 > 0$, and that $h(\cdot)$ satisfies (A1), we have

$$\max_{t \le T} \|h(\bar{x}(\omega, t))\| \le K_2(\omega) < \infty \tag{36}$$

$$E \max_{t \le T} \|h(\bar{x}(\omega, t))\| \le K_3 < \infty, \tag{37}$$

for some $K_2(\omega)$ and K_3. Now, since $h(\bar{x}(\cdot, \cdot))$ and $h(\bar{x}(\omega, \cdot))$ are measurable functions, the weak convergences (31) and (32) imply that

$$q_\omega^n(t) \to 0$$
$$E\, q_\omega^n(t) \to 0 \qquad (38)$$

pointwise in t, $t \leq T$. In addition, by virtue of (36) and (37) and the boundedness of $k(\cdot)$, both $q_\omega^n(t)$ and $E\, q_\omega^n(t)$ satisfy Lipschitz conditions in t; i.e.,

$$\|q_\omega^n(t) - q_\omega^n(t - h)\| \leq K_4(\omega)h. \qquad (39)$$

Thus, for $t \leq T$, both $q_\omega^n(t)$ and $E\, q_\omega^n(t)$ tend to zero uniformly in t. Equation (36) and (37) imply

$$\bar{\epsilon}_\omega^{\,n} = \sup_{t \leq T} \|q_\omega^n(t)\|/t < K_8(\omega) < \infty \qquad \text{w.p.1.}$$
$$\bar{\epsilon}^n = \sup_{t \leq T} E\|q_\omega^n(t)\|/t < \infty \qquad \text{w.p.1.} \qquad (40)$$

Now, applying (A1), (39) and (40) to (34) yields, for some K_5,

$$\|\delta x^n(\omega, t)\| \leq \int_0^t K_5\|\delta x^n(\omega, s)\|ds + \bar{\epsilon}_\omega^{\,n} t,$$

which implies

$$\|\delta x^n(\omega, t)\| \leq K_6\, \bar{\epsilon}_\omega^{\,n} t \qquad (41)$$

and

$$E \max_{t \leq T} \leq K_7\, \bar{\epsilon}^n \qquad (42)$$

for some K_6 and K_7. Thus (29) and (30) hold.

Let

$$T_\omega = \liminf_n T_\omega^n. \tag{43}$$

The remainder of the proof is exactly as in Theorem 1 and will be omitted.

The proof of the Corollary follows easily from the proof of the theorem, and will not be given.

Corollary. Assume the conditions of Theorem 4, except that $f_0(x, u)$ may take any sign and $z(\cdot, \cdot)$ may not be continuous, and the terminal condition is the set of expectations (4). Also the terminal time is non-random and bounded by some $\hat{T}$. Then, if there is one control satisfying (4) and with a terminal time bounded by $\hat{T}$, there is an optimal control satisfying (4) and with terminal time bounded by $\hat{T}$.

References.

1. H. J. Kushner "On the Stochastic Maximum Principle with 'Average' Constraints", to appear in J. Math. Anal. Appl.

2. L. S. Pontriagin, et. al., The Mathematical Theory of Optimal Processes, Interscience, 1962, translated by K. N. Trirogoff and L. W. Neustadt.

3. E. B. Lee, L. Marcus, "Optimal Control for Nonlinear Processes", Arch. Rational Mech. Anal. 8 (1961) pp. 36-58.

4. E. Roxin, "The Existence of Optimal Controls", Mich. Math. J., 9 (1962) pp. 109-119.

5. H. J.Kushner, "On Stochastic Extremum Problems; Calculus", to appear, J. Math. Anal. Appl.

6. H. J. Kushner, "On the Stochastic Maximum Principle; Fixed Time of Control", to appear, J. Math. Anal. Appl.

7. S. L. Doob, "Stochastic Processes", Wiley, (1953).

8. M. Loeve, Probability Theory, Van Nostrand, 3rd Edition (1963).

9. T. Wazewski, "Sur une Condition d'Existence des Fonctions Implicites Measurables", Bull. de l'Acad. Polon. des Sci., Serie des Sciences, Math. Astr. et Phys., 9 (1961) pp. 861-863.